

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: February 11, 2004, 10:44:14 ; Search time 33 Seconds  
(without alignments)  
1433.951 Million cell updates/sec

Title: US-09-441-723-1  
Perfect score: 1185  
Sequence: 1 MGPLPRTVELFDVLSYPSW.....AHLLGKWMGPIPPAVNARL 226

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pdb.p:  
2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pdb.p:  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW PUB.pdb.p:  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.p:  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pdb.p:  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS PUBCOMB.pdb.p:  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pdb.p:  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.p:  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pdb.p:  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pdb.p:  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pdb.p:  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pdb.p:  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pdb.p:  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pdb.p:  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pdb.p:  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pdb.p:  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pdb.p:  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.p:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		ID	Description
		Match	Length		
1	1185	100.0	240	9	US-09-876-889-349
2	326.5	27.6	107	11	Sequence 349, App
3	113	9.5	241	15	Sequence 5132, Ap
4	92.5	7.8	255	15	Sequence 11189, A
5	91	7.7	581	12	Sequence 8857, Ap
6	89.5	7.6	962	12	Sequence 2888, Ap
7	89	7.5	545	12	Sequence 23533, A
8	89	7.5	592	12	Sequence 20, Appl
9	89	7.5	592	12	Sequence 2, Appl1
10	89	7.5	592	12	Sequence 4, Appl1
11	89	7.5	592	12	Sequence 64, Appl
12	89	7.5	609	12	Sequence 67, Appl
13	89	7.5	609	12	Sequence 6, Appl1
14	89	7.5	751	12	Sequence 63, Appl
15	86.5	7.5	395	10	Sequence 59, Appl
					US-09-738-626-4728

SUMMARIES

Query Match	100.0%;	Score	1185;	DB	9;	Length	240;
Best Local Similarity	100.0%;	Pred. NO.	2.9e-120;				
Matches	226;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MGPLPRTVELFDVLSYPSWLGFEILCRYQNIWNINQLRPSLITGIMKDSGNKPPGLLP	60				
Db	15	MGPLPRTVELFDVLSYPSWLGFEILCRYQNIWNINQLRPSLITGIMKDSGNKPPGLLP	74				
Qy	61	RKGLYMANDLKLRRHILQIPHFDPFLSVMLEKGSLSAMRFLTAVNLEHPMELEKASRE	120				
Db	75	RKGLYMANDLKLRRHILQIPHFDPFLSVMLEKGSLSAMRFLTAVNLEHPMELEKASRE	134				
Qy	121	LWWRVSRNEDITEPQSILAAAEKAGMSAQOGLLEKIATPKVKNQIKETTEACRYGA	180				
Db	135	LWWRVSRNEDITEPQSILAAAEKAGMSAQOGLLEKIATPKVKNQIKETTEACRYGA	194				

ALIGNMENTS

RESULT 1  
US-09-876-889-349  
; Sequence 349, Application US/09876889  
; Patent No. US20020076715A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN  
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.466C3  
; CURRENT APPLICATION NUMBER: US/09/876,889  
; CURRENT FILING DATE: 2001-06-06  
; NUMBER OF SEQ ID NOS: 353  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 349  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-876-889-349

Sequence 6230, Ap  
Sequence 10510, A  
Sequence 14132, A  
Sequence 14693, A  
Sequence 10, Appl  
Sequence 286, App  
Sequence 92, Appl  
Sequence 8637, Ap  
Sequence 12, Appl  
Sequence 1864, Ap  
Sequence 6324, Ap  
Sequence 45, Appl  
Sequence 113, App  
Sequence 23358, A  
Sequence 2, Appl1  
Sequence 18, Appl  
Sequence 18917, A  
Sequence 14491, A  
Sequence 11264, A  
Sequence 988, App  
Sequence 21509, A  
Sequence 115, App  
Sequence 1901, Ap  
Sequence 9749, Ap  
Sequence 12296, A  
Sequence 497, App  
Sequence 432, App  
Sequence 10, Appl  
Sequence 53, Appl  
Sequence 2, Appl1



APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 CURRENT FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 2888  
 LENGTH: 581  
 TYPE: PRF  
 ORGANISM: Thermotoga maritima  
 US-10-369-493-2888  
  
 Query Match 7.7%; Score 91; DB 12; Length 581;  
 Best Local Similarity 19.0%; Pred. No. 0.84; Mismatches 35; Indels 38; Gaps 6;  
 Matches 35; Conservative 44;  
  
 QY 28 RYQNIW-----NINLQRLPSLITGIMKDSGNKPPGL-----LPRKGLYMAN 68  
 DB 343 RPNVWFVSYDGKWLKNDINLDFQPGKLYAIVGETGGKSTLMSLINGLYIPQKGNIFID 402  
  
 QY 69 DLKLRHHLQIP-----IHFPKDFL-----SVMLEKSGLSAMRFLTAVNLEHP-EMLE 115  
 DB 403 EIPLEYNLKLVRKQIAAQPQVLLFSGTILNIRLFDESIPERVLEALKRVHALDIIE 462  
  
 QY 116 KASRELWMRVWSNEDITEPQSILAAAEKAGM-----SAEQAGLEKIATPKVNOLK 169  
 DB 463 RLPGGVYIEVERGTGTLASGERQLIALARAVLFDKIFILDEATSNVDVITETKIOALE 522  
  
 QY 170 ETTE 173  
 DB 523 ELSK 526  
  
 RESULT 6  
 US-10-369-493-23533  
 Sequence 23533, Application US/10369493  
 Publication No. US20030233675A1  
 GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 CURRENT FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 23533  
 LENGTH: 962  
 TYPE: PRF  
 ORGANISM: Escherichia coli  
 US-10-369-493-23533  
  
 Query Match 7.6%; Score 89.5; DB 12; Length 962;  
 Best Local Similarity 26.3%; Pred. No. 2.6; Mismatches 30; Indels 23; Gaps 3;  
 Matches 30; Conservative 13;  
  
 QY 126 WSNEDITEPQSILAAAEKAGMSEQA-----QGLEKIATPKVNQ 167  
 DB 731 WCRNKDVVDKQSVIFEKAGNSTDSALAAVFPVTGYDEYTSAYSSLLGQIVQVPWFYNQ 790  
  
 QY 168 LKETEACRYAGFGLPITVAHVDGQTHLFGSDRMELLAAHLGCEKWMGPIPPA 221

DB 791 LR--TEEQLGAVFAFPMSVGRQWGWGFLQSNKQ---PSFLWERYKAPFPTA 839  
  
 RESULT 7  
 US-10-413-943-20  
 Sequence 20, Application US/10413943  
 Publication No. US20040006784A1  
 GENERAL INFORMATION:  
 APPLICANT: Mourad, George S.  
 TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
 TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
 FILE REFERENCE: PRF-07898  
 CURRENT APPLICATION NUMBER: US/10/413,943  
 CURRENT FILING DATE: 2003-04-15  
 NUMBER OF SEQ ID NOS: 69  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 20  
 LENGTH: 545  
 TYPE: PRF  
 ORGANISM: Arabidopsis thaliana  
 US-10-413-943-20  
  
 Query Match 7.5%; Score 89; DB 12; Length 545;  
 Best Local Similarity 22.8%; Pred. No. 1.3; Mismatches 49; Indels 80; Gaps 11;  
 Matches 49; Conservative 29;  
  
 QY 55 PPGL-LPRKGLYMANDLKLRLHHLQIPHFPPKDFLSVM-----LEKSGLS-AMRFLT-- 104  
 DB 5 PPKLPLPR-----LKVSPNSLQ-----YPAGYLGAVPERTNEAENGSGIAEAMEYLTNI 52  
  
 QY 105 -----AVNLEHP-EMLEKASRELWMRVWSNEDITE-----POSILAA 141  
 DB 53 LSTKVYDIAIESPLQLAKLKRGLVRYLKRDLQPVFSFKLRGAYNMVVKLPADQLAK 112  
  
 QY 142 AEKAGMSAQAG-----LLEKIATPKVNQKETTETACRYGAFGLPITVA 188  
 DB 113 GVTCSSAGNHAQGVALSASKLGCTAVIVMPVTTPTEIKWQAVENL----- 156  
  
 QY 189 HVDGQTHLFGSDRMELLAAHL--LGEKWMGPIPP 220  
 DB 157 ---GATVVLFGSDYDQAQAHAKIRABEEGLTTRIPP 188  
  
 RESULT 8  
 US-10-413-943-2  
 Sequence 2, Application US/10413943  
 Publication No. US20040006784A1  
 GENERAL INFORMATION:  
 APPLICANT: Mourad, George S.  
 TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
 TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
 FILE REFERENCE: PRF-07898  
 CURRENT APPLICATION NUMBER: US/10/413,943  
 CURRENT FILING DATE: 2003-04-15  
 NUMBER OF SEQ ID NOS: 69  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 2  
 LENGTH: 592  
 TYPE: PRF  
 ORGANISM: Arabidopsis thaliana  
 US-10-413-943-2  
  
 Query Match 7.5%; Score 89; DB 12; Length 592;  
 Best Local Similarity 22.8%; Pred. No. 1.4; Mismatches 49; Indels 80; Gaps 11;  
 Matches 49; Conservative 29;  
  
 QY 55 PPGL-LPRKGLYMANDLKLRLHHLQIPHFPPKDFLSVM-----LEKSGLS-AMRFLT-- 104  
 DB 52 PPKLPLPR-----LKVSPNSLQ-----YPAGYLGAVPERTNEAENGSGIAEAMEYLTNI 99  
  
 QY 105 -----AVNLEHP-EMLEKASRELWMRVWSNEDITE-----POSILAA 141

Db 100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLPQVFSFKLRGAYNNMMVKLPADQLAK 159  
QY 142 AEKAGMSAEQAQG-----LLEKIATPKVNQKLTETEAACRYGAFGLPITVA 188  
Db 160 GVICSSAGNHAQGVALSASKLGCTAVIMVPTTPEIKWQAVENL-----203  
QY 189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPPIP 220  
Db 204 ---GATVVLFGDSYDQQAQAHAKIRAEIEGLTFIPP 235

## RESULT 9

US-10-413-943-4  
; Sequence 4, Application US/10413943  
; Publication No. US20040006784A1  
; GENERAL INFORMATION:  
; APPLICANT: Mourad, George S,  
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
; FILE REFERENCE: PRF-07898  
; CURRENT APPLICATION NUMBER: US/10/413,943  
; CURRENT FILING DATE: 2003-04-15  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 592  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
US-10-413-943-4

Query Match 7.5%; Score 89; DB 12; Length 592;  
Best Local Similarity 22.8%; Pred. No. 1.4; Mismatches 57; Indels 80; Gaps 11;  
Matches 49; Conservative 29;

QY 55 PPGL-LPRKGLYMANDLKLRRHLQIPHFDPFLSV-----LEKGSLS-AMRFLT-- 104  
Db 52 PPKLPLPR-----LKVSPNSLQ---YPAGYLGAVPERTNEAENGSIAMBYLTNI 99  
QY 105 -----AVNLEHP-EMLEKASRELMMRVMSRNEDITE-----POSILAA 141  
Db 100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLPQVFSFKLRGAYNNMMVKLPADQLAK 159  
QY 142 AEKAGMSAEQAQG-----LLEKIATPKVNQKLTETEAACRYGAFGLPITVA 188  
Db 160 GVICSSAGNHAQGVALSASKLGCTAVIMVPTTPEIKWQAVENL-----203  
QY 189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPPIP 220  
Db 204 ---GATVVLFGDSYDQQAQAHAKIRAEIEGLTFIPP 235

## RESULT 10

US-10-413-943-64  
; Sequence 64, Application US/10413943  
; Publication No. US20040006784A1  
; GENERAL INFORMATION:  
; APPLICANT: Mourad, George S,  
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
; FILE REFERENCE: PRF-07898  
; CURRENT APPLICATION NUMBER: US/10/413,943  
; CURRENT FILING DATE: 2003-04-15  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 64  
; LENGTH: 592  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
US-10-413-943-64

Query Match 7.5%; Score 89; DB 12; Length 592;  
Best Local Similarity 22.8%; Pred. No. 1.4; Mismatches 57; Indels 80; Gaps 11;  
Matches 49; Conservative 29;

QY 55 PPGL-LPRKGLYMANDLKLRRHLQIPHFDPFLSV-----LEKGSLS-AMRFLT-- 104  
Db 52 PPKLPLPR-----LKVSPNSLQ---YPAGYLGAVPERTNEAENGSIAMBYLTNI 99  
QY 105 -----AVNLEHP-EMLEKASRELMMRVMSRNEDITE-----POSILAA 141  
Db 100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLPQVFSFKLRGAYNNMMVKLPADQLAK 159  
QY 142 AEKAGMSAEQAQG-----LLEKIATPKVNQKLTETEAACRYGAFGLPITVA 188  
Db 160 GVICSSAGNHAQGVALSASKLGCTAVIMVPTTPEIKWQAVENL-----203  
QY 189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPPIP 220  
Db 204 ---GATVVLFGDSYDQQAQAHAKIRAEIEGLTFIPP 235

## RESULT 11

US-10-413-943-67  
; Sequence 67, Application US/10413943  
; Publication No. US20040006784A1  
; GENERAL INFORMATION:  
; APPLICANT: Mourad, George S,  
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
; FILE REFERENCE: PRF-07898  
; CURRENT APPLICATION NUMBER: US/10/413,943  
; CURRENT FILING DATE: 2003-04-15  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67  
; LENGTH: 592  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
US-10-413-943-67

Query Match 7.5%; Score 89; DB 12; Length 592;  
Best Local Similarity 22.8%; Pred. No. 1.4; Mismatches 57; Indels 80; Gaps 11;  
Matches 49; Conservative 29;

QY 55 PPGL-LPRKGLYMANDLKLRRHLQIPHFDPFLSV-----LEKGSLS-AMRFLT-- 104  
Db 52 PPKLPLPR-----LKVSPNSLQ---YPAGYLGAVPERTNEAENGSIAMBYLTNI 99  
QY 105 -----AVNLEHP-EMLEKASRELMMRVMSRNEDITE-----POSILAA 141  
Db 100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLPQVFSFKLRGAYNNMMVKLPADQLAK 159  
QY 142 AEKAGMSAEQAQG-----LLEKIATPKVNQKLTETEAACRYGAFGLPITVA 188  
Db 160 GVICSSAGNHAQGVALSASKLGCTAVIMVPTTPEIKWQAVENL-----203  
QY 189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPPIP 220  
Db 204 ---GATVVLFGDSYDQQAQAHAKIRAEIEGLTFIPP 235

## RESULT 12

US-10-413-943-6  
; Sequence 6, Application US/10413943  
; Publication No. US20040006784A1  
; GENERAL INFORMATION:  
; APPLICANT: Mourad, George S,  
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms  
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase  
; FILE REFERENCE: PRF-07898  
; CURRENT APPLICATION NUMBER: US/10/413,943  
; CURRENT FILING DATE: 2003-04-15  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 609

```

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-413-943-6

Query Match
Best Local Similarity 7.5%; Score 89; DB 12; Length 609;
Matches 49; Conservative 29; Mismatches 57; Indels 80; Gaps 11;

Qy 55 PPGL-LPRKGLYMANDLKLRLHLLQIPHFPPKDFLSV-----LEKGSLS-AMRFLT-- 104
Db 69 PPKLPLPR-----LKVSPNSLQ-----YPAGYLGAVPERTNEAENGSAEAMEYLTNI 116
Qy 105 -----AVNLEHP-EMLEKASRELWMRVMSRNEDEITE-----POSILAA 141
Db 117 LSTKVYDIAIESPLQAKLSKRLGVRMYLKREDLQPVFSFKLRGAYNMVVKLPADQLAK 176
Qy 142 AEKAGMSAEQAQG-----LLEKIATPKVKNQKLTETEAACRYGAFGLPTIVA 188
Db 177 GVICSSAGNHAQGVALSASLKGCTAVIVMPVTTTPEIKQAVENL----- 220
Qy 189 HVDGQTHMLFGSDRMELLALH-----LGEKWMGPIPP 220
Db 221 ---GATVVLFGDSYDQQAQAHAKIRAESEGLTFIPP 252

RESULT 13
US-10-413-943-63
; Sequence 63, Application US/10413943
; Publication No. US20040006784A1
; GENERAL INFORMATION:
; APPLICANT: Mourad, George S,
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase
; FILE REFERENCE: PRF-07898
; CURRENT APPLICATION NUMBER: US/10/413.943
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-413-943-63

Query Match
Best Local Similarity 7.5%; Score 89; DB 12; Length 609;
Matches 49; Conservative 29; Mismatches 57; Indels 80; Gaps 11;

Qy 55 PPGL-LPRKGLYMANDLKLRLHLLQIPHFPPKDFLSV-----LEKGSLS-AMRFLT-- 104
Db 69 PPKLPLPR-----LKVSPNSLQ-----YPAGYLGAVPERTNEAENGSAEAMEYLTNI 116
Qy 105 -----AVNLEHP-EMLEKASRELWMRVMSRNEDEITE-----POSILAA 141
Db 117 LSTKVYDIAIESPLQAKLSKRLGVRMYLKREDLQPVFSFKLRGAYNMVVKLPADQLAK 176
Qy 142 AEKAGMSAEQAQG-----LLEKIATPKVKNQKLTETEAACRYGAFGLPTIVA 188
Db 177 GVICSSAGNHAQGVALSASLKGCTAVIVMPVTTTPEIKQAVENL----- 220
Qy 189 HVDGQTHMLFGSDRMELLALH-----LGEKWMGPIPP 220
Db 221 ---GATVVLFGDSYDQQAQAHAKIRAESEGLTFIPP 252

RESULT 14
US-10-413-943-59
; Sequence 59, Application US/10413943
; Publication No. US20040006784A1
; GENERAL INFORMATION:
; APPLICANT: Mourad, George S,
; TITLE OF INVENTION: Methods and Compositions for Producing Plants and Microorganisms
; TITLE OF INVENTION: that Express Feedback Insensitive Threonine Dehydratase/Deaminase

```

```

; FILE REFERENCE: PRF-07898
; CURRENT APPLICATION NUMBER: US/10/413.943
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-413-943-59

Query Match
Best Local Similarity 7.5%; Score 89; DB 12; Length 751;
Matches 49; Conservative 29; Mismatches 57; Indels 80; Gaps 11;

Qy 55 PPGL-LPRKGLYMANDLKLRLHLLQIPHFPPKDFLSV-----LEKGSLS-AMRFLT-- 104
Db 60 PPKLPLPR-----LKVSPNSLQ-----YPAGYLGAVPERTNEAENGSAEAMEYLTNI 107
Qy 105 -----AVNLEHP-EMLEKASRELWMRVMSRNEDEITE-----POSILAA 141
Db 108 LSTKVYDIAIESPLQAKLSKRLGVRMYLKREDLQPVFSFKLRGAYNMVVKLPADQLAK 167
Qy 142 AEKAGMSAEQAQG-----LLEKIATPKVKNQKLTETEAACRYGAFGLPTIVA 188
Db 168 GVICSSAGNHAQGVALSASLKGCTAVIVMPVTTTPEIKQAVENL----- 211
Qy 189 HVDGQTHMLFGSDRMELLALH-----LGEKWMGPIPP 220
Db 212 ---GATVVLFGDSYDQQAQAHAKIRAESEGLTFIPP 243

RESULT 15
US-09-738-626-4728
; Sequence 4728, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4728
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4728

Query Match
Best Local Similarity 7.5%; Score 88.5; DB 10; Length 395;
Matches 38; Conservative 24; Mismatches 65; Indels 35; Gaps 7;

Qy 84 PKDFLSVMLEKGSLSAMRFLTAVNLEHP-EMLEKASRELWMRVMSRNEDEITEPOSILAAAE 143
Db 206 PEALMAFLMEKIQLTKYLFTHP-KHPE-----QVMSPDYGDIGPEAYANATL 253

```

Qy 144 KAGMSAEQACGLLEKJATPKVK-NOLKETTEAACRYGAFGLPITVAHVD-----GOT 194  
Db 254 VCAKDLDEVAGATEKSYTSEKMKALIRARDGHCRFPGCCVPASKCQVDHIIIPWABGGPT 313  
Qy 195 -----HMLFGSDRMELLALLGE-KWMGPJP-PAV 222  
Db 314 AAWNLOLLCQRHNMTDGRFTADANGLAEIRWIGPMDVPAV 355

Search completed: February 11, 2004, 10:49:59  
Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 10:39:59 ; Search time 21 Seconds  
(without alignments)  
455.345 Million cell updates/sec

Title: US-09-441-723-1  
Perfect score: 1185  
Sequence: 1 MGPLPRIVELFDVLPVSM.....AHLGERKMGPIPPAVNARL 226

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1185	100.0	226	3	US-08-978-174-1
2	874	73.8	226	3	US-08-978-174-3
3	196	16.5	203	4	US-09-252-991A-17282
4	87	7.3	1150	4	US-09-252-991A-24671
5	80.5	6.8	339	2	US-08-855-714-3
6	78.5	6.6	353	4	US-09-252-991A-16824
7	77	6.5	659	4	US-09-198-452A-432
8	76.5	6.5	317	2	US-09-066-075-2
9	76.5	6.5	317	3	US-08-518-615A-2
10	76.5	6.5	317	3	US-08-951-889-2
11	76.5	6.5	317	3	US-09-472-857-2
12	76.5	6.5	497	1	US-08-075-193-4
13	76.5	6.5	497	2	US-08-564-090A-4
14	76.5	6.5	497	5	PCT-US94-06698-4
15	76	6.4	588	4	US-09-601-777-2
16	75.5	6.4	1398	1	US-08-750-532-9
17	75.5	6.4	1398	3	US-08-894-818B-8
18	75.5	6.4	1398	4	US-09-445-472-6
19	75	6.3	289	4	US-09-134-001C-4583
20	74	6.2	339	4	US-09-489-847-350
21	74	6.2	554	4	US-09-489-847-352
22	74	6.2	884	4	US-09-741-150-4
23	74	6.2	2474	4	US-08-265-967C-3
24	74	6.2	2474	4	US-08-305-790B-4
25	73.5	6.2	404	4	US-09-328-352-6854
26	73.5	6.2	523	2	US-08-473-553A-3
27	73.5	6.2	869	1	US-08-188-582-32

Query Match 100.0%; Score 1185; DB 3; Length 226;

28	73.5	6.2	869	1	US-08-646-715-32	Sequence 32, Appl
29	73.5	6.2	980	2	US-08-473-553A-6	Sequence 6, Appli
30	73.5	6.2	985	2	US-08-473-553A-2	Sequence 2, Appli
31	73	6.2	733	4	US-09-345-473B-42	Sequence 42, Appl
32	73	6.2	948	1	US-08-698-551-14	Sequence 14, Appl
33	73	6.2	948	2	US-08-602-228-14	Sequence 14, Appl
34	73	6.2	948	2	US-08-533-901B-14	Sequence 14, Appl
35	73	6.2	948	2	US-08-839-032A-14	Sequence 14, Appl
36	73	6.2	948	2	US-08-839-031A-14	Sequence 14, Appl
37	73	6.2	948	4	US-09-185-258C-14	Sequence 14, Appl
38	73	6.2	948	5	PCT-US95-12724-14	Sequence 14, Appl
39	72.5	6.1	249	4	US-09-107-532A-5923	Sequence 523, Ap
40	72	6.1	803	4	US-09-252-991A-19492	Sequence 19492, A
41	71.5	6.0	461	4	US-09-346-408-8	Sequence 8, Appli
42	71.5	6.0	532	3	US-09-181-336-15	Sequence 15, Appl
43	71.5	6.0	543	2	US-08-922-170B-10	Sequence 10, Appl
44	71.5	6.0	543	3	US-09-071-739B-2	Sequence 2, Appli
45	71.5	6.0	543	3	US-09-181-336-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-978-174-1  
; Sequence 1, Application US/08978174  
; Patent No. 6030809  
; GENERAL INFORMATION:  
; APPLICANT: Shah, Purvi  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: NEW GLUTATHIONE-S-TRANSFERASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fatseq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,174  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0430 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 226 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BLADTUT04  
; CLONE: 1554593  
; US-08-978-174-1





PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 24671  
LENGTH: 1150  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24671

Query Match 7.3%; Score 87; DB 4; Length 1150;  
Best Local Similarity 26.4%; Pred. No. 1.3;  
Matches 46; Conservative 27; Mismatches 59; Indels 42; Gaps 11;  
QY 84 PKDPLSVMLEKGSLSAMRFLTAVN-LEHPMELEKASRELWMRYWSRNE-----DITEPSQI 138  
DB 581 PEDWLC-----DGTGYDFNQSLSLQHDPRGRPIRLMQRVSGRPEAFLDVYQARQL 635  
QY 139 LAAAEKAGMSAEQAQGLLEKIATPKVKNQKETTEAACRYGAFGL-----PI--TVAHVDPG 192  
DB 636 VLAGSLAGDLENLAQGLL-RVARADLAS--RDLTGGIRRALFOLLARFPVYRTYAGACG 692  
QY 193 QT-----HMLFGSDRMELLAHLGKWMG-----PIPPAVNARL 226  
DB 693 RSVQDREVFYAAEAAREDLDEADR-AVLDEL--ERWLGQQLRELPFGPLREL 743

RESULT 5  
US-08-855-714-3  
Sequence 3, Application US/08855714  
Patent No. 5939075  
GENERAL INFORMATION:  
APPLICANT: Hough, Huo-Shu H.  
APPLICANT: Warren, Richard L.  
TITLE OF INVENTION: MUTANTS OF BRUCELLA MELITENSIS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John Moran, Esq.  
STREET: HQ USAMRDC, Dept. of Army, Fort Detrick  
CITY: Frederick  
STATE: MD  
COUNTRY: US  
ZIP: 21702-5012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/855,714  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/334,129  
FILING DATE: 04-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenna  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: 08/143,692  
TELEPHONE: (301) 619-2065  
TELEFAX: (301) 619-7714  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 339 amino acids  
TYPE: amino acid  
STRANDEDNESS: both  
TOPOLOGY: unknown  
MOLECULE TYPE:  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: BRUCELLA MELITENSIS  
US-08-855-714-3

Query Match 6.8%; Score 80.5; DB 2; Length 339;  
Best Local Similarity 21.7%; Pred. No. 1.1;  
Matches 65; Conservative 25; Mismatches 96; Indels 113; Gaps 14;  
QY 3 PLPRTVELFYDVL-----PYSWLGFELCRYQIWININLQRLPSLITGIMKDSGNKPPGLL 59  
DB 47 PLPTSSPMNLKMCQSAPPTSLKRRLLFCPPPHWK-----SLRTASQKSSSFTKAALK 99  
QY 60 PRK-GLYMANDIKLRRHHLIQIPIHPFKDFLSVMLEKGS-----SAMRFLTAVNLEH 110  
DB 100 PRPGSMMWTRKSSPRASALGRAWH-PQDRRLGYDGGQVRLASLDLDTQACNAFAIN-KA 157  
QY 111 PEMLE---KASRELWMRYWSRNEDETEPOSILAAAEKAG-----SVIAARDRSNVAIFDLAENVHKGILATST 202  
DB 158 PALLEGVFVFEFV-----MSAQOAGLLEKIA-----TPKVNOLKETTE 173  
QY 147 -----MSAQOAGLLEKIA-----TPKVNOLKETTE 173  
DB 203 VPAATISVQTAAEAARTAAEKLLHALDYGVGLGLEFFVLKDGTLTLLANEFAPRVHNS-GHWTE 261  
QY 174 AACRYGAF-----GLPI--TVAHVDCQTHMLFGSD-----RMELLAHLGK 213  
DB 262 AACATISQFEQHIRAVAGLPLGNTDRHSDCMENLIGDIEKVPAILCEKXNAVILHLYGKK 320

RESULT 6  
US-09-252-991A-16824  
Sequence 16824, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16824  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16824

Query Match 6.6%; Score 78.5; DB 4; Length 353;  
Best Local Similarity 26.1%; Pred. No. 2;  
Matches 62; Conservative 26; Mismatches 79; Indels 71; Gaps 15;  
QY 33 WVINLQRLPSLITGIMKDSGNK-----PPGLLPKGLYMA-----NDLKLRRHHLIQIPI 81  
DB 134 WTALLVRKDSPIRSLAELKGRKVAATKGTDPYFLRLSLHSLVGLDKNDLRIV--HLQHPD 191  
QY 82 HFPKDFLSVMLEKGSLSAMRFLTAVNLEHPM-----LEKASRELWMR-----VMSRNE 130  
DB 192 G-----RVALEKQVDAWAGL-----DPHMAASELOAGSRLLYNLFNSYGVNLVRE 239  
QY 131 DITE--PQ---SILAAAEKA-----GMSAEQAQGLLEKIATP--KVNOLKETTEAACRY 178  
DB 240 DFAERHPQLIRQVLAAYEQARHWVIGHDPDEAQLAAEAGLPLEVARLQISRTD----- 293  
QY 179 GAFGLPI-----TVAHVDCQTHMLF-----GSDRMELLAHLGKWK-----NGPIPPA 221  
DB 294 --FSQPLGAEQVAALKAAAPILADERLVRPGVDQKVVDDELIAPOWAAAEVIGGVPLA 349  
RESULT 7  
US-09-198-452A-432  
Sequence 432, Application US/09198452A  
Patent No. 6559294

```

; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 432
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...659
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-432

Query Match      6.5%; Score 77; DB 4; Length 659;
Best Local Similarity 21.3%; Pred. No. 7.8;
Matches 52; Conservative 36; Mismatches 98; Indels 58; Gaps 8;

QY 9 ELFYDVL-----PYSWLGPEILCRYQNI-----WNINLQLRPSLITGI 47
DB 371 ERFYEVNLHPDLHSQKEREIEFLGLSNTITFENVSGYQEDKHILKNLSFTLHKGEALGI 430
QY 48 MKDSGNKPP-----GLLRPKGLYANDLKLRRHHLQIPHPKDFLSVMLEKGSLSAMRFLT 104
DB 431 VGPTSGKTLVKLLPR-----LYEVSQKILDISLPITEY-----NKGSL-----RNHI 475
QY 105 AVNLEHP-----EMLEKASRELWVRVMSNRNEDITEPQSILAAAEKAGWSA 149
DB 476 ACVLQNPFLFYDTVWNNLTGCKOMBEZAVLEAKRAYADEFLKLPKGVHVSLESGKNL 535
QY 150 EQAQGLLEKIATPKVKNO-----LKETTEAACRYGAFGLPITVAHVDDGQTHMLFGSDRMEL 205
DB 536 SGGQOORLAIALRALLKNASILDLDEATSDALDAISENYIKNIIGELKGQCTQIIIAHKLTT 595
QY 206 LAHL 209
DB 596 LEHV 599

RESULT 8
US-09-066-075-2
; Sequence 2, Application US/09066075
; Patent No. 5925749
; GENERAL INFORMATION:
; APPLICANT: Mathur, E., et al.
; TITLE OF INVENTION: Carboxymethyl Cellulase from Thermotoga Maritima
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,075
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 331400-20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 331400-20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-09-066-075-2

Query Match      6.5%; Score 76.5; DB 2; Length 317;
Best Local Similarity 21.1%; Pred. No. 2.9;
Matches 43; Conservative 22; Mismatches 50; Indels 89; Gaps 10;

QY 76 HLOIPIH-----PP-----KDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEK 116
DB 49 HVRPIRSTHAYAPPPYKIMDRFFKRVDEINGALKKG-----LAVAINIHVYELMN 102
QY 117 ASRE-----LWVRVMSNRNED-----ITEPQS----- 137
DB 103 DPEEHKFRFLWKQIADRYKDYPTLFFELNBPNGHNLTPKWNELLEALKVIRSIDK 162
QY 138 -----ILAAAEKAGWSAQGLLEKIATPKVKNOKLTETEAACRYGAFGLPITVAHVDDGQ 193
DB 163 KHTIIIGTAEWGGSIA-----LEKLSVPKWE-----KNSIVTHYINPPEF----- 203
QY 194 THMLFGSDRMELLAHLLGEKWMGP 217
DB 204 THQ--GAWEVGESEKNLGRKMGSP 225

RESULT 9
US-08-518-615A-2
; Sequence 2, Application US/08518615A
; Patent No. 5962258
; GENERAL INFORMATION:
; APPLICANT: Mathur, E., et al.
; TITLE OF INVENTION: Carboxymethyl Cellulase from Thermotoga Maritima
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/518,615A
; FILING DATE: August 23, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 331400-20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

US-08-951-889-2

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

[illegible]

Db 49 HVRIPRSTHAYAPPPYKMDRPFKRVDEVINGALKRG-----LAVAINIHYYEELMN 102  
QY 117 ASRE-----LWVRVWSNED-----ITEPQS----- 137  
Db 103 DEEHKERFLAKQIADRYKDYPTLPPPEILNEPHGNLTPEKWNELLEALKVIRSIDK 162  
QY 138 -----ILAAAKAGMAEQAGLEKIAIPKVNQKLEKTEAAACRYGAGLPITVAHVQGO 193  
Db 163 KHTIIIGTAEWGGISA-----LEKLSVPKWE---KNSIVTIHYNPFEE----- 203  
QY 194 THMLFGSDRMELLHLGLSKWGP 217  
Db 204 THQ--GAWEVGESEKWLGRKWGP 225

RESULT 12  
US-08-075-193-4  
; Sequence 4, Application US/08075193  
; Patent No. 5547868  
; GENERAL INFORMATION:  
; APPLICANT: MILLER, WALTER L.  
; APPLICANT: HARIKRISHNA, JENNIFER A.  
; APPLICANT: BLACK, STEPHEN M.  
; TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/075,193  
FILING DATE: 09-JUN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: NEELEY PH.D., RICHARD L.  
REGISTRATION NUMBER: 30,092  
REFERENCE/DOCKET NUMBER: UCAL-236/0005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-494-7622  
TELEFAX: 415-857-0663  
TELEX: 380816 COOLEY PA  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-075-193-4

Query Match 6.5%; Score 76.5; DB 1; Length 497;  
Best Local Similarity 25.4%; Pred. No. 5.8;  
Matches 44; Conservative 29; Mismatches 69; Indels 31; Gaps 9;

QY 14 VLSPSYWLGFELCRYQNIWNINQLRPSLTITGIMKDSGNKPPGLLPKRG-LYMANDLKL 72  
Db 195 LITPPEHLEALLCQRTDITKAAL-----GVLRQSRVKTWLVGRGPIQVAFITKE 246  
QY 73 LRHHLQI-----PIHFPKDFL-----SVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121  
Db 247 LREMIQLPGARPILDVDFGLQDKIKEVPRPKRLTELLLTAT--EKPGPAEAARQAS 304  
QY 122 WVRVWSRNEDITEPOSILAAAEKAGMSAQGLEKIAIPKVNQKLEKTEA 174  
Db 305 ASRAGL-RFPSPQVLPSPD-----GRRAGV--RLAVTELEG-VDEATRA 348

RESULT 13  
US-08-564-090A-4  
; Sequence 4, Application US/08564090A  
; Patent No. 5939318  
; Patent No. 5939318 5741703  
; GENERAL INFORMATION:  
; APPLICANT: MILLER, WALTER L.  
; APPLICANT: HARIKRISHNA, JENNIFER A.  
; APPLICANT: BLACK, STEPHEN M.  
; TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: COOLEY GODWARD LLP  
; STREET: FIVE PALO ALTO SQUARE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/564,090A  
FILING DATE: 02/05/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD L. NEELEY, PH.D.  
REGISTRATION NUMBER: 30,092  
REFERENCE/DOCKET NUMBER: UCAL-236/01US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5000  
TELEFAX: 415-857-0663  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-564-090A-4

Query Match 6.5%; Score 76.5; DB 2; Length 497;  
Best Local Similarity 25.4%; Pred. No. 5.8;  
Matches 44; Conservative 29; Mismatches 69; Indels 31; Gaps 9;

QY 14 VLSPSYWLGFELCRYQNIWNINQLRPSLTITGIMKDSGNKPPGLLPKRG-LYMANDLKL 72  
Db 195 LITPPEHLEALLCQRTDITKAAL-----GVLRQSRVKTWLVGRGPIQVAFITKE 246  
QY 73 LRHHLQI-----PIHFPKDFL-----SVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121  
Db 247 LREMIQLPGARPILDVDFGLQDKIKEVPRPKRLTELLLTAT--EKPGPAEAARQAS 304  
QY 122 WVRVWSRNEDITEPOSILAAAEKAGMSAQGLEKIAIPKVNQKLEKTEA 174  
Db 305 ASRAGL-RFPSPQVLPSPD-----GRRAGV--RLAVTRLEG-VDEATRA 348

RESULT 14  
PCT-US94-06698-4  
; Sequence 4, Application PC/TUS9406698  
; GENERAL INFORMATION:  
; APPLICANT: MILLER, WALTER L.  
; APPLICANT: HARIKRISHNA, JENNIFER A.  
; APPLICANT: BLACK, STEPHEN M.  
; TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBBINS, BERLINER & CARSON

Qy 58 LLPRKGLYMANDLKLRRHHLQPIHF-PKDFLSVMLEKSGLSAMRFLTAVNLEHPEMLEK 116  
Db 73 ALPRPA--QAQDVVDLDFTTQEPHLVSPFSLVTIDANLATDPRELIILGSPKLRTLAR 130  
Qy 117 ASRELWRYVMSRNE--ITEPQSILAAAEKAGMSAEQAQGLLEKIAATPKVKNQKETTEA 174  
Db 131 GLSPAYLRFGGTKTDFLIFDPKKESTFEERSYQSQVNQDI----- 171  
Qy 175 ACRYGA 180  
Db 172 -CKYGS 176

Search completed: February 11, 2004, 10:45:28  
Job time : 22 secs

STREET: 201 NORTH FIGUEROA STREET  
CITY: LOS ANGELES  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/06698  
FILING DATE: FILED HERewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: BERLINER, ROBERT  
REGISTRATION NUMBER: 20,121  
REFERENCE/DOCKET NUMBER: 5555-224-C1  
TELEPHONE: 213-977-1001  
TELEFAX: 213-977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-06698-4

Query Match 6.5%; Score 76.5; DB 5; Length 497;  
Best Local Similarity 25.4%; Pred. No. 5.8;  
Matches 44; Conservative 29; Mismatches 69; Indels 31; Gaps 9;

Qy 14 VLSPSYWLGFELCRYQNIWNINQLRPSLITGIMDSGNKPPGLPRKG-LYMANDLKL 72  
Db 195 LITPPEHLEALLCQRTDITKAAL-----GVLRQSRVKTVMVGRRGFLQVAFIKE 246  
Qy 73 LRHHLQI-----PIHFPKDFL-----SVMLEKSGLSAMRFLTAVNLEHPEMLEKASREL 121  
Db 247 LREMIQLPGARPIIDPVDVDFGLQDKIEVPRPRKRLTELLRLTAT--EKPGPAEAFQAS 304  
Qy 122 WMRVWSRNEITEPQSILAAAEKAGMSAEQAQGLLEKIAATPKVKNQKETTEA 174  
Db 305 ASRAWGL-RFRSPQQVLPSPD-----GRRAGV--RLAVTRLEG-VDEATRA 348

RESULT 15  
US-09-601-777-2  
Sequence 2, Application US/09601777  
Patent No. 6461848  
GENERAL INFORMATION:  
APPLICANT: Nakajima, Motowo  
APPLICANT: Funakubo, Minako  
TITLE OF INVENTION: Human heparanase polypeptide and cDNA  
FILE REFERENCE: 30384A  
CURRENT APPLICATION NUMBER: US/09/601,777  
CURRENT FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 588  
TYPE: PRT  
ORGANISM: Human  
US-09-601-777-2

Query Match 6.4%; Score 76; DB 4; Length 588;  
Best Local Similarity 19.4%; Pred. No. 8.5;  
Matches 36; Conservative 27; Mismatches 71; Indels 52; Gaps 6;

Qy 14 VLSPSYWLGFELCRYQNIWNINQLRPSLITGIMDSGNKPPGLPRKG-LYMANDLKL 72  
Db 24 ISAPFKW-----VWVISWRGQPGPKMLRSKALPPPLMLLLGLPLGSLPG 72